Basis Representation Theorem - Alternate Proof

James Philip Rowell

January 4, 2019 (v01)

Basis Representation Theorem

Let b be a positive integer greater than 1.

For every positive integer n there is a unique sequence of integers $d_0, d_1, d_2, \ldots, d_k$ such that:

$$n = d_k b^k + d_{k-1} b^{k-1} + \cdots + d_2 b^2 + d_1 b^1 + d_0 b^0,$$

where $0 \leq d_i < b$ for all i in $\{0, 1, 2, \ldots, k\}$ and $d_k \neq 0$.

Definition: n is represented in base-b by the string of base-b-digits $(d_k d_{k-1} \cdots d_2 d_1 d_0)_b$

The paper "Counting[*]" proves the above theorem by induction, but suggests that it could also be proven by generalizing a technique[†] used to calculate the digits of a number for a given base. The following proof uses that approach, involving repeated divisions of n by the base b, the remainders of which end up being the base-b digits of n.

Lemma

Let b be an integer where $b \neq 0$ and $c_0, c_1, c_2, \ldots, c_n$ be a sequence of integers, then:

$$(((\ldots(((c_0)b+c_1)b+c_2)b+\ldots+c_{n-2})b+c_{n-1})b+c_n) = c_0 b^n + c_1 b^{n-1} + c_2 b^{n-2} + \ldots + c_{n-2} b^2 + c_{n-1} b^1 + c_n b^0$$

Proof of Lemma by Induction

Base case:

When $n = 1$ we have $(c_0)b + c_1 = c_0 b^1 + c_1 b^0$, and also note that the lemma holds for $n = 0$ since $(c_0) = c_0 b^0$.

Induction step:

Assume the lemma is true for $n = k$ and prove it true for $n = k + 1$.

$$\begin{aligned}
&((((\ldots(((c_0)b + c_1)b + c_2)b + \ldots + c_{k-2})b + c_{k-1})b + c_k)b + c_{k+1}) \\
&= ((c_0 b^k + c_1 b^{k-1} + c_2 b^{k-2} + \ldots + c_{k-2} b^2 + c_{k-1} b^1 + c_k b^0)b + c_{k+1}) \\
&= c_0 b^{k+1} + c_1 b^k + c_2 b^{k-1} + \ldots + c_{k-2} b^3 + c_{k-1} b^2 + c_k b^1 + c_{k+1} b^0
\end{aligned}$$

QED

As a reminder, a statement of the "Euclidean Division Theorem" follows,

[*] Also written by James Rowell.

[†] Exercise 2-iii from the paper "Counting".

Euclidean Division Theorem

For all integers a and b such that $b > 0$, there exist *unique* integers q and r such that:

$$a = qb + r \text{ such that } 0 \leq r < b$$

Definition: In the above equation:

a is the *dividend*	("the number being divided")
b is the *divisor*	("the number doing the dividing")
q is the *quotient*	("from Latin *quotiens* 'how many times' b goes into a")
r is the *remainder*	("what's left over (if anything) after the division")

Proof of Basis Representation Theorem

Let b be a positive integer greater than 1 and let n be a positive integer.

Dividing n by b we get non-negative integers q_1 and d_0 such that,

$$n = q_1 b + d_0; \text{ where, } 0 \leq d_0 < b.$$

If $q_1 \neq 0$ we continue this process by dividing b into q_1 to get integers q_2 and d_1 such that,

$$q_1 = q_2 b + d_1; \text{ where, } 0 \leq d_1 < b,$$

As long as the new quotient (i.e., q_2) is non-zero, we continue this process until we get a quotient, say $q_{k+1} = 0$, as follows,

$$\begin{aligned} q_2 &= q_3 b + d_2; \text{ where, } 0 \leq d_2 < b \\ q_3 &= q_4 b + d_3; \text{ where, } 0 \leq d_3 < b \\ \ldots \\ q_{k-1} &= q_k b + d_{k-1}; \text{ where, } 0 \leq d_{k-1} < b \\ q_k &= q_{k+1} b + d_k; \text{ where, } 0 \leq d_k < b \end{aligned}$$

There will be an integer k for which $q_{k+1} = 0$ but $q_k \neq 0$, because for all q_i in the above list of equations we have,

$$\begin{aligned} q_i &= q_{i+1} b + d_i \\ &\geq q_{i+1} b + 0 \\ &\geq 2q_{i+1} \\ &> q_{i+1} \end{aligned}$$

Letting $q_0 = n$, then the above strict-inequality leads us to conclude that,

$$q_0 > q_1 > q_2 > q_3 > \ldots > q_k > q_{k+1}.$$

Since no quotients are negative then the sequence must terminate with $q_{k+1} = 0$ for some $k \geq 0$.* We must note that in this case, $d_k \neq 0$, since if it were then $q_k = 0$, which can't be true otherwise the process would have stopped one step earlier.

*As an interesting aside, $k = \lfloor log_b(n) \rfloor + 1$.

Back-substituting each expression for q_{i+1} into the expression for q_i, starting with q_{k+1},

$$\begin{aligned}
q_{k+1} &= 0, \\
q_k &= 0 \cdot b + d_k, \\
q_{k-1} &= (d_k)b + d_{k-1}, \\
q_{k-2} &= ((d_k)b + d_{k-1})b + d_{k-2}, \\
\ldots,
\end{aligned}$$

hence,

$$n = (((\ldots(((d_k)b + d_{k-1})b + d_{k-2})b + \ldots d_2)b + d_1)b + d_0)$$

By an application of our lemma (noting the change of indices: $d_k = c_0, d_{k-1} = c_1, \ldots,$ $d_1 = c_{k-1}, d_0 = c_k$), then we can conclude that:

$$n = d_k b^k + d_{k-1} b^{k-1} + \cdots + d_2 b^2 + d_1 b^1 + d_0 b^0.$$

Furthermore $0 \leq d_i < b$ for all i in $\{0, 1, 2, \ldots, k\}$ and $d_k \neq 0$.

Furthermore the "Euclidean Division Theorem" guarantees that each sequence of integers $d_0, d_1, d_2, \ldots, d_k$ is unique because each q_i and d_i resulting from each division is unique.

QED